



ARE DIGITAL LAWS MAKING OR BREAKING DIGITAL LIBRARIES?

Benjamin White

First Published in 2011, t.LXXXII, 1-4. Archives et Bibliothèques de Belgique / Archief – En Bibliotheekwezen in België

CC BY 3.0

Abstract: *Information law in the digital world presents libraries with many legal challenges. Whether this is trying to enjoy the same types of rights in the digital era that they enjoyed in the analogue world, being able to embark on the mass digitisation of their 20th century collections under appropriate new laws, or simply continuing to offer internet access in the face of graduated response or three strikes enforcement legislation. This article seeks to give a high level overview of many of the legal developments (copyright, contract law, legal deposit, and IPR enforcement) that have affected libraries in the last five to ten years. The views expressed in the article are the author's alone.*

Introduction

Tim Berners-Lee's completion of the world's first web browser (WorldWideWeb) on Christmas Day 1990 irrevocably changed the norms of the information world. From production to access of information, and through to information sharing how the world's more than 2 billion plus internet citizens engage with digital information has transformed our intellectual landscape forever. As storehouses of knowledge this change presents libraries, and therefore this generation of librarians and information professionals, with huge and exciting challenges as well as opportunities. In order to engage with the digital environment and to perform their core role of information provision, libraries require a number of core criteria to be fulfilled. At a fundamental level hardware, software, and the availability of stable internet provision are prerequisites in order to provide basic services to users of a modern day library. Moreover the larger research and national libraries are having to engage with the digital environment in new and interesting, but sometimes legally and politically challenging ways, in order to become the fully-fledged providers of digital collections that users and society now expect of them.

This article will explore at an overview level some of the legal and political challenges that have affected the library community since the start of the digital revolution. Given the unquestionable central role that digital information now has in society, the implications of how well the sector itself is able to engage and deal with these legal challenges presents in many peoples' eyes a defining moment in the history of the centuries old profession of librarianship.

Libraries on the Net and the Political Landscape

A literature survey of the period 1994 -1995 shows that libraries were quick to grasp the life-changing impact that the web was going to have on society, and in particular how people would want to access scholarly information as well as other forms of knowledge in the future. The first UK listing of 100 websites published in Issue 1, October 1994 of 'Internet' (Emap Publishers), which at the time it proclaims was "the most extensive ever published in a magazine", was well populated with websites from the library and cultural sector. It is interesting to note that the

number of libraries and museums online at that time appear to be far more numerous than certain categories one would assume to have been better represented such as finance, government or sport. It lists for example website offerings from the British Library, the Louvre, the Vatican Library, as well as a very early library-type website offering full text works as varied as Shakespeare and the latest scripts from Star Trek. Another early example of a library website listed was a 1994 online exhibition from the Library of Congress on the history of the Vatican complete with pictures, text and hyperlinks, supplementing the library's real world exhibition.¹

As libraries began to digitise parts of their collection during the 1990s through programmes like the Electronic Libraries Programme (eLib) in the UK, Memoria Hispanica in Spain, Global Info in Germany, or the Digital Library Initiative (DLI) in America, one of the inevitable issues to arise would be copyright law. Because of its complexities, combined with extensive pre-20th century collections of many research libraries, a significant proportion of the large-scale digitisation projects in the 1990s, and indeed still to this day, have focussed solely on out-of-copyright works. Given the time, complexity and expense of rights clearance this represents a pragmatic approach given the numerous unavoidable costs that any large scale digitisation project will face as a matter of course, such as project management, the cost of digitisation itself, and then the sustainable provision of a technical infrastructure to preserve and host the digitised content.

While the focus of many library digitisation projects remains pre-20th century material, the societal and political debate around what to digitise started to change rapidly in 2004 when Google entered into partnership with five high-profile libraries (Oxford, University of Michigan, Harvard, Stanford and the New York Public Library) in what was known then as the Google Print Library Project, and now simply Google Books. The aim of the project was to digitise a mixture of out of copyright material, as well as crucially, in-copyright material which Google argued was allowable under the US legal doctrine of 'fair use'.

This became a decisive turning point in the political and legal discourse around mass digitisation particularly in the European Union. Strong concerns were raised about the Google library projects within the European political establishment, witnessed by a letter from six EU heads of state to the Commission² in April 2005 calling for the establishment of a European virtual library. The catalyst for this was in fact an editorial from the then head of the French National Library Jean-Noël Jeanneney that had appeared in *Le Monde*³ earlier in 2005. Jeanneney argued that the Google Books project represented not only a threat to the cultural diversity of the web, being a predominantly English language project, but that it should also be countered by a free multilingual European Digital Library based upon a transparent search hierarchy. This led to the establishment of the i2010 Digital Libraries programme in 2006 which comprised various working groups and High Level Expert Groups. The groups were tasked with looking at funding issues, as well as the main copyright and legal barriers to mass digitisation, such as orphan works⁴, licensing mechanisms for out of print in-copyright works, as well as digital preservation. Following on from this Europeana, an EU funded portal to cultural works available from over 1500 European institutions' websites, and the very thing Jeanneney called for, was launched to much fanfare in November 2008. This was accompanied the same month by the establishment of the ARROW project to build a European wide database of rights information, including orphan works and public domain works. In 2011 a report on mass digitisation by an independently appointed committee of experts (*Comité des Sages*)⁵ was published and most recently in May 2011 we see perhaps a legal culmination of much of this activity in the form of the publication by the European Commission of a draft directive on orphan works.

¹ The online exhibition in part is still viewable on You Tube. <http://www.youtube.com/watch?v=YF8jWUR1yAE> [Accessed 22/07/2011]

² http://ec.europa.eu/information_society/activities/digital_libraries/background/index_en.html [Accessed 22/07/2011]

³ Jeanneney, J-N. 2005. *Quand Google défie l'Europe*. *Le Monde*. 23/01/2005.

⁴ Orphan works are in-copyright works whose rightsholder cannot be identified or located.

⁵ De Decker, J. M Lévy, E Niggemann. 2011. *The New Renaissance: Report of the Comité des Sages. Reflection Group on Bringing Europe's Cultural Heritage Online*.

http://ec.europa.eu/information_society/activities/digital_libraries/doc/reflection_group/final-report-cdS3.pdf [Accessed 22/07/2011]

Beyond the EU at a global level, libraries have also been involved in agenda setting, and have been the focus of much political discourse. Libraries have been one of many civil society groups that have played an integral part in the current movement to reform the norm setting agenda of the World Intellectual Property Organisation. This movement, which became an officially adopted workplan in 2007, is known as the Development Agenda. Since the discussions on the Broadcast Treaty at WIPO broke down in 2007/8, discussions on limitations and exceptions – a prime focus of the Development Agenda – have been an integral part of deliberations at WIPO’s Standing Committee on Copyright and Related Rights (SCCR). As part of this new agenda a number of studies commissioned by the SCCR on the issues of access to knowledge by the visually impaired, libraries and education were published between 2006 and 2009⁶. Linked to this in 2008 a draft text for a Treaty for the Blind, Visually Impaired and Other Reading Disabled Persons (TVI) was published by the World Blind Union (WBU) who had been active by this point for many years at WIPO. Crucially for libraries off the back of the ground-breaking work done by the WBU, in April 2011 the International Federation of Library Associations (IFLA) published a draft Treaty on Copyright Exceptions and Limitations for Libraries and Archives.⁷

While libraries have been the focus of much policy debate, this should not be confused with libraries having political influence per se. Other than at the WIPO level, which was the result of high levels of political engagement by IFLA and other bodies like Electronic Information for Libraries (eIFL), and the American Library Association (ALA), the influencing capabilities and level of engagement by library groups in the sphere of information law has been somewhat limited in recent years. Rather than the current focus on digitisation issues by the Commission (mainly DG Information Society and Media) being the result of library effort as such, it is more to do with Google, the perceived Americanization of online culture, and a political realization of the economic and social benefits of digitisation that has led to the political developments that we see. This lack of concerted effort by libraries to influence the development of the legislative agenda, particularly in Europe since the passing of the Information Society Directive in 2001, is perhaps in part connected to the increasingly difficult legal and financial position that libraries find themselves in.

Whither Library and Research Exceptions?

It is perhaps no exaggeration to say that whilst without copyright law libraries would be a pale shadow of the institutions they are today, in the digital era copyright in its current form is proving to be a significant barrier to knowledge sharing in the context of scholarly information. The commodification of information – the consistent expansion of intellectual property into areas traditionally considered to be public domain – has been a constant factor affecting the users of copyright, particularly over the 20th century (Elkin-Koren and Netanel, 2002). In addition to this, the very fact the technology allows you to reuse copyright material so easily, and the clear social and economic benefits that ensue, has made people rethink whether copyright law as it is currently formulated is nurturing or hindering the development of society and whether it is contributing to public welfare as it should. This is particularly the case in North America where a small group of academics (Lawrence Lessig, James Boyle, Michael Geist etc), many of them associated with the Creative Commons movement, have been particularly vocal beyond the hallways of academia on issues such as copyright and what is referred to as the “enclosure” of the public domain.

Copyright law which only lasted 14 years when first implemented in Great Britain in 1710 with the *Statute of Anne*, now in the main lasts for life plus 70 years after the death of the author. Over the 20th century intellectual property rights have been extended greatly to cover databases, the publishers of old manuscripts, performers, software producers, broadcasters and now through

⁶ Crews, K. 2008. *Study on Copyright Limitations and Exceptions for Libraries and Archives*. Standing Committee on Copyright and Related Rights. World Intellectual Property Organisation. Sullivan, J. 2006. *Study on Copyright Limitations and Exceptions for the Visually Impaired*. Standing Committee on Copyright and Related Rights. World Intellectual Property Organisation. Xalabarder, R. 2009. *Study on Copyright Limitations and Exceptions for educational activities in North America, Europe, Caucasus, Central Asia and Israel*. Standing Committee on Copyright and Related Rights. World Intellectual Property Organisation.

⁷ <http://www.ifla.org/en/publications/draft-treaty-on-copyright-exceptions-and-limitations-for-libraries-and-archives> [Accessed 22/07/2011]

contract law even a couple of words can be protected. In addition to the seemingly inexorable expansion of intellectual property rights, users of copyright works such as libraries face a double legislative whammy in that neither existing limitations and exceptions are being uniformly updated, nor are new laws being introduced quickly enough to facilitate new knowledge sharing activities made possible by technological advancement.

One example of this failure to update existing laws is in the field of digital preservation – one of the core functions of legal deposit libraries. While the European Union has harmonised the ‘exclusive rights’ granted to creators of copyright works, the limitations and exceptions or permitted acts in the European Union have not been harmonised. From the perspective of libraries this creates a confusing picture where by dint of national legislature some countries have digital exceptions, while others essentially do not. For example to all intents and purposes either Italy or Lithuania allow digital preservation (Crews, 2008). Even where digital copies can be made for preservation purposes, legal limitations are idiosyncratically placed around the capability of a library to preserve items within their own collections. For example in Latvia and Sweden computer programmes are excluded from digital library preservation and in the UK neither sound recordings nor film can be preserved lawfully. Similarly the number of preservation copies that can be made is often limited by law. For example it is limited to one in the UK, Greece, Hungary, Poland, and Latvia, three in Slovenia, and eight by case law in Germany. Despite the fact that the activity of copying by mirror servers and format shifting are basic principles and necessities of digital preservation, these legal restrictions make digital preservation effectively unlawful in a not insignificant number of European countries.

Research exceptions also follow this idiosyncratic trend. In Europe research exceptions either exist in their own right or are subsumed within private copying exceptions. Despite the Information Society Directive allowing “reproductions on any medium made by a natural person for private use” (Article 5(2)b), as we see with digital preservation exceptions, the rules around the copies that can be made for research purposes vary dramatically from country to country within the single market. Many countries allow research copying in all and any media, while in some countries like France this is limited to photocopying. What can be copied also varies from country to country. For example Belgium, Denmark and Estonia limit copying to published works, the UK excludes sound and film from research copying, Hungary excludes copying of public performances and Latvia and Slovenia exclude databases. (Xalabarder, 2009)

Given that preservation and research exceptions are arguably the core limitations and exceptions for research libraries - existing to preserve cultural and information heritage and then allow access to it - the lack of norm setting in this area says much about the nature of the global copyright debate. Whereas we have seen numerous recent changes to legislation benefiting rightsholder groups such as term extension⁸ and the creation of numerous new intellectual properties over the 20th century, there has been relatively little focus on aspects of intellectual property rights that promote education and learning. Given that access to knowledge underpins education, research and therefore economic growth downstream, many in the library and education world would argue that the need for a more balanced policy agenda around intellectual property laws is well overdue.

Another extremely important issue impacting on the ability of library users to use copyright works in the digital age is in the area of contract law. Until the 1990s most access to copyright goods was simply regulated by copyright law, which by recognising the unique nature and societal significance of information, aims to balance both public and private interests. Now however much access to electronic information is regulated by contract law. Whereas copyright law, despite being subject to much lobbying does have at its heart a notion of balance, in order to protect the interests of creators and users, contract law is not subject to any such checks and balances. In regards to knowledge goods at best it is an unequal power play between owner and customer in a process of negotiation aimed at acquiring access to the copyright works in question. But the reality of many digital contracts is that they are in fact not negotiated at all. Whereas in theory an

⁸ In September 2011, the EU voted to extend the term for performers and sound recordings from 50 to 70 years
<http://www.bbc.co.uk/news/technology-14829373>

end user can negotiate an End User Licence Agreement (EULA) to access digital content, or a library a digital database, in practice most digital licences are offered on a 'take it or leave it' basis. The *laissez-faire* concept of freedom to contract is an important part of free market economies, but the lack of checks and balances around contracts in the sphere of information law also contributes to a shrinking of the public space where knowledge can be freely shared and built upon – namely the public domain. A number of predominantly American copyright academics have referred to this general trend of increasing barriers to access to knowledge, of which increasing regulation by contract represents only one facet, as the “second enclosure movement.”⁹

Copyright laws, particularly in developed countries, as outlined above very often have limitations and exceptions to facilitate the activities that take place in libraries – including typically preservation, research copying, copying for the visually impaired. However contract law as a private contract between two legal entities simply sets out a relationship as envisaged by the licensor of the copyright work, whether that be a piece of music from iTunes, a CD-ROM, or a scientific database costing millions of Euros. This represents a fundamental issue and something that is affecting the traditional balance between the public and private interest. The little empirical evidence that exists on this issue shows that the majority of contracts offered to libraries are more restrictive than national copyright law would allow libraries or their users to do.¹⁰ While legal theory regarding the law of statute holds that a balance between the interests of the creator of a copyright work and the user are very much needed in order to stimulate further innovation and creativity, thus ensuring that the monopoly granted to the creator does not have an adverse effect on the flow of knowledge in society, no such legal interventionist thinking exists in the field of contract law for libraries, or their users, in the vast majority of WIPO member states.¹¹

In 2008 the British Library undertook a study of 100 contracts. Referring to this it is very clear that for publicly funded and therefore publicly accountable bodies like libraries, the trend to provide copyright goods under contract is fundamentally affecting the norms of information access. While acknowledging the complexities of trying to simply and easily compare and benchmark contractual provisions to statute based limitations and exceptions, the study nevertheless attempts to illustrate how typical library exceptions provided by law are reflected, or not as the case may be, in the contracts being offered to research libraries. In the case of preservation, the study shows that only 23 contracts permitted archiving, 19 expressly excluded it and 58 of the contracts were silent. Similarly only two contracts out of 100 allowed for copying for or by the visually impaired, and only 14 allowed interlibrary loan – 4 expressly excluded it, while the remainder were silent. However contracts are of course not necessarily always a one way street. While in the main the study shows a worrying undermining of limitations and exceptions, it is important to point out that many of the licences allowed for copying by individuals for research or private study purposes, to a more generous level than would be allowed for under an appropriate research or private copying exception in law.

So what does this mean for libraries and their users? The reliance upon contract law for access to digital content is not only undermining the permitted acts expressly granted by the legislature, with all that implies for the role of copyright law itself in the digital world, it is also increasing the cost to access information itself. Even ignoring the above inflation increases that are well known in the journal market, it can be argued that the cost of purchasing information has risen significantly as the rights to use information that come with digital objects are less than they are in the analogue world. If further reuse of knowledge goods is covered by copyright law then as part of the price what is effectively bundled in are certain statutorily backed 'reusage rights'. However it appears that these 'reusage rights' are not uniformly part of the deal when it comes to digital copyright goods that are licensed. Not only are they not uniformly part of the deal, but there are quantitatively less of them as the British Library study of the 100 contracts demonstrates, as even common library related exceptions like digital preservation or copying for the visually impaired are

⁹ Boyle, J. 2008. *The Public Domain: Enclosing the Commons of the Mind*. Yale University Press.

¹⁰ *An analysis of 100 Contracts offered to the British Library*

<http://pressandpolicy.bl.uk/ImageLibrary/detail.aspx?MediaDetailsID=691> [Accessed 29/07/2011]

¹¹ The exceptions to this are Ireland, Portugal and Belgium who do not allow for contracts to over-ride limitations and exceptions in statute. Also the EU Database Directive does not allow contracts to over-ride limitations and exceptions for database rights.

usually not provided for in a contract. Given that the cost of purchasing licensed material goes up year on year for libraries, and yet the 'reusage rights' are declining, it can therefore be argued that in the digital age the costs upstream to libraries are not only going up for accessing information in the first place, but also downstream, as less can now be done with the copyright goods purchased.

Transaction costs and the burden of managing information legally have also increased. Exceptions such as 'fair use' and 'fair dealing' or indeed library exceptions often referred to as 'library privilege' allow individuals as well as librarians to navigate access to knowledge within relatively well understood parameters. Activities such as research copying, inter-library loan, preservation copying, copying for the visually impaired as well as other exceptions such as judicial and parliamentary copying, are based on decades old interpretations of the law and are something that forms part of any professional course on librarianship. In addition to this in many countries there is broad agreement between libraries and publisher groups as to how to practically apply the law in the case of personal or research copying.¹² In such a context the transaction costs of navigating and intermediating limitations and exceptions are relatively low as the terms of access are well understood as linked to long tried and tested concepts in law.

This contrasts strikingly with the norms of digital information, and in particular scholarly information that is purchased or licensed in the form of databases, CD-ROMs and other digital media. A large British university's annual acquisitions budget may be as high as £20 million per annum, of which an increasingly large amount is spent on digital products. However in addition to exceptions being undermined by rising costs and diminishing usage rights, libraries are finding that the transaction costs of managing each individual digital object lawfully have increased as each contract or licence's terms and conditions vary. The question therefore naturally arises: given the vast amounts of digital scholarly information being purchased - estimated to be over \$27 billion globally in 2007¹³ – how can object by object lawful access to knowledge be managed in a cost effective manner by institutions such as libraries?

The position of a library patron is even more complex. The reality being that in the majority of instances an end user in a library is actually of course not even aware of what they can or can't do with a digital database or an e-book as the contract is probably held by the library's contracts officer and will most likely incorporate a confidentiality clause preventing the end user having access to it.

With this background, library groups at a national level as well as a supranational level are seeking a revision of copyright law to ensure that limitations and exceptions are no longer overridden by contract law. Without this libraries argue that further norm setting agendas regarding limitations and exceptions are pointless, as any change to statute will simply be negated by contract. They also argue that because of the supremacy of contract law, that innovation, creativity and economic development itself is being threatened by the creation of *de facto* information monopolies through the imposition of private contracts on users of copyright goods. Although this issue has attracted relatively little attention at a policy level, the interplay of copyright and contract is an issue that cuts to the heart of the copyright agenda in the digital age as without addressing it, the very role of copyright law in the digital age is called into question. Whereas governments have traditionally chosen to intervene in how information is created, monetised and distributed in order to protect the interests of creators and their users for varied private and public interest purposes, if this is no longer seen as a role for government, the repercussions of this will be likely be severe for knowledge sharing and innovation in the future.

Mass Digitisation, the Law and Libraries

Since Google announced its decision to partner with some of the biggest brands in the education world in 2004 – Oxford, Harvard, New York Public Library, Stanford and Michigan – the issues of mass digitisation and copyright law have never been far from the minds of law makers

¹² Clark, C. 1990. *Photocopying from books and journals: a guide for all users of copyright works*. British Copyright Council.

¹³ 2008 Library Market Size, Share and Forecast Report. Outsell.

internationally. Since 2005 the American courts have also been dealing with a class action law suit between first of all Google and the American Authors Guild, and subsequently Google and the Association of American Publishers in regards to Google's digitisation of books in US libraries. This was followed in April 2010 by another class action infringement suit against Google launched by the American Society of Media Photographers along with the Graphic Artists Guild, Picture Archive Counsel of America, North American Nature Photography Association, and the Professional Photographers of America. More recently in September 2011, and again linked to Google's library projects, the American Authors Guild, and a handful of individual writers, also filed an action against five American universities and the digital repository, the Hathi Trust. The universities were sharing on their linked campus networks orphan works¹⁴ that had been digitised by Google, asserting they had the right to do so under "fair use".¹⁵

In essence what is at stake both sides of the Atlantic in the context of mass digitisation is how to deal with two key issues – one being orphan works, the other being the mass clearance of rights – in a way that pragmatically facilitates digitisation but does not undermine the exclusive rights and legitimate interests of copyright holders. The proponents of mass digitisation including in particular DG Information Society and Media argue that the societal, educational and economic benefits that will come from facilitating the digitisation of Europe's great libraries and archives will be manifold. Certainly it is perhaps not an understatement to say that the digital opening up of these great storehouses of knowledge in Europe, America and beyond will have an immeasurable but undeniably positive effect on learning, our own view of the world, and unleash greater creativity and innovation in turn.

Currently orphan works create a barrier to any mass digitisation project, whether carried out by a library, archive or a private company such as Google. An orphan work is an in-copyright item whose creator cannot be identified or located. As outlined by an EU study (Vuopala, 2008) it affects all forms of copyright works, creating a significant financial and legal barrier to digitisation. A couple of examples in the Vuopala report highlight the costs and time spent in clearing rights that often are simply not possible to clear. For example the National Archives in the UK spent two years and an estimated £35,000 in trying to clear 1,114 wills of which less than half were ever cleared. Another case study from Austria shows that despite the clear public interest in the online access of 200,000 German language dissertations from 1925 – 1988 the transaction costs of rights clearance for this type of non-commercial material with inevitably high levels of orphans would be an estimated 20-50 times higher than the cost of digitisation itself, and they therefore remain digitised but inaccessible online.

The rate of orphan works will vary depending on the nature of the collection. A survey from the Association des Cinémathèques Européennes in 2010 found that 21% of works in European film archives are orphan works. Other studies from the US¹⁶ and the UK¹⁷ of published monographs show that rates of at least 30% and higher are not uncommon. Similarly certain sound recording collections also have significant level of orphan works. For example, according to the EU study (Vuopala, 2008) the African Writers Club collection, that forms part of the British Library's Archival Sound Recording project, contains orphan rates of 64% for literary works and 85% for the performances.

Prior to the current debate on orphan works a number of countries had already introduced limited provisions for the use of orphan works in national legislation. Since 1985 Article 77 of the Canadian Copyright Act has provided for a licence based solution that is granted by the Canadian Copyright Board upon being presented with evidence that "reasonable efforts to locate the owner of the copyright" have been made. A licence based solution to orphan works was also introduced into Japanese copyright law in 1970. Article 67 which applies only to works that have been made public, allows for the granting of a compulsory licence from the Commissioner of the Agency for

¹⁴ A number of titles were found very quickly not to be orphan works by the Authors Guild leading to a reevaluation of the workflows by the universities involved. <http://www.lib.umich.edu/news/u-m-library-statement-orphan-works-project>

¹⁵ *Legal Action on College Book Plan*. <http://www.bbc.co.uk/news/technology-14896567> [Accessed 18/09/2011]

¹⁶ Troll Covey, D. 2005. *Acquiring Copyright Permission to Digitize and Provide Open Access to Books*. Digital Library Federation.

¹⁷ Stratton, B. 2011

Cultural Affairs (*Bunkachō*), upon due diligence and the payment of a royalty fee which the agency believes to be “corresponding to an ordinary rate of royalty.”

Understandably at a policy level an awakening of interest in this issue globally was prompted by the fact that mass digitisation became a technological reality around the year 2000. Despite this, legislative change has been slow coming. Out of 27 EU member states so far only two have successfully implemented a provision to deal with the issue of orphan works, with arguably Finland not requiring a change to their copyright act, though usage is limited to certain types of works only. Denmark amended its copyright law (Art 50.2) in 2008 allowing collecting societies to license orphan works. In Hungary an amendment to the 1999 copyright law was adopted by Parliament in 2009 allowing the Hungarian Patent Office to offer non-exclusive licences to use orphan works. The conditions for the re-licensing by government of orphan works are time limited to a usage period of 5 years, and do not allow sub-licensing or the creation of derivative works, so to what extent this provision proves useful has yet to be seen.

Both the UK and US legislatures have attempted to pass bills to allow the use of orphan works though both attempts failed for different reasons. In the UK as part of the Digital Economy Bill introduced to parliament in 2010, a provision allowing for the government to license orphan works to third parties was proposed by the Labour government. However as part of a UK specific parliamentary process known as ‘wash-up’, where normal parliamentary practice is abandoned, the orphan works provision was dropped in horse-trading between the Labour and Conservative parties due to opposition from a well-organised group of independent photographers. In the US, opposition to the 2008 Orphan Works Bill also came from photographers as well as the National Writers Union. Most surprisingly perhaps, opposition also came from Lawrence Lessig of the Creative Commons movement, who writing in the *New York Times* declared his opposition to the bill saying that it was an “amazingly onerous and inefficient change” to American copyright law.¹⁸ In the face of this mounting opposition the bill that had started in the Senate lost all momentum, and in fact never reached the House of Representatives.

The orphan works issue is of great importance to libraries as it is an integral part of the mass digitisation legal jigsaw puzzle. This is perhaps best demonstrated by the fact that it is part of the draft international library treaty currently being discussed at WIPO. Article 13 of the draft treaty states that in order to reduce financial and legal liabilities for libraries “It shall be permissible for libraries and archives to reproduce and otherwise use a work, and material protected by related rights, for which the author or other rightholder cannot be identified or located after reasonable inquiry, including by making the work available to the public.”

At an EU level after many years of meetings, working groups, and discussions of ‘soft law’ a proposed directive for orphan works was finally published in May 2011. While library groups have welcomed the draft proposal as a first step in addressing the issue of mass digitisation, they have also pointed out that the directive has a number of serious flaws that will act as a barrier to European culture being made comprehensively available online in any meaningful sense.¹⁹ The proposed directive as it currently stands (September 2011) excludes unpublished materials that are an important part of any large library or archive’s collection, as well as sound recordings, and stand alone artistic works. Legal deposit libraries, museums and a number of other cultural institutions are also excluded from using films, and a country of origin principle has been introduced whereby only material from the EU can be treated as orphan in line with the directive. The requirement to do a diligent search for each and every rightsholder under Art 3(1) also presents a major problem for libraries and other organisations as this may require a physical inspection of each item to look for embedded works such as photographs, illustrations or poems. Given that initial estimates from the British Library²⁰ and the Royal Library of Denmark suggest that illustrations appear in nearly half of all books, item by item physical inspection of books and

¹⁸ Lessig, L. 20/05/2008. *Little Orphan Artworks*. *New York Times*.
<http://www.nytimes.com/2008/05/20/opinion/20lessig.html?ex=1369022400&en=af6d685002b2942f&ei=5124&partner=permalink&exprod=permalink>

¹⁹ http://www.eblida.org/uploads/Statement%20on%20Orphan%20Works%20Directive%20proposal_EBLIDA%20-%20LIBER-ENCES-27May2011.pdf [Accessed 18/09/2011]

²⁰ 46.7% of records from the British National Bibliography contained an illustration field.

monographs presents an insurmountable barrier to mass digitisation. By requiring a diligent search for each embedded work, the draft directive also curiously ignores the *raison d'être* of the Commission funded ARROW database that is intended to avoid manual diligent search by facilitating automated diligent search through the provision of a comprehensive network of rights databases. A recently published British Library/ARROW report (Stratton, 2011) found that 43% of a small sample of 140 books published between the 1870s and 2010 were orphan works. Whereas on average the study spent 4 hours per book establishing its copyright status and looking for rightsholders, the ARROW system took 5 minutes per record showing the importance of using well-constructed databases to facilitate mass digitisation.

Despite the high levels of political muscle that have been put into the issue of orphan works, mass digitisation itself is unlikely to ever to be comprised solely of orphan works. The large political focus on orphan works therefore is perhaps therefore best understood as a less controversial starting point in addressing the issue of how to encourage mass digitisation of 20th century copyright works in general. Copyright works that could be digitised can be divided into four categories – public domain and therefore copyright has expired, in-copyright and yet orphan, in-copyright but out-of-commerce, and currently in-commerce and in-copyright. The issue of out of commerce works has recently gained traction in DG Internal Market led discussions and has opened up a new frontier in the digital policy debate.

Large numbers of in-copyright works that are of out-of print or out of commerce sit in libraries and archives and return no financial or reputational reward to their creator. The ARROW study (Stratton, 2011), while an extremely small sample indeed, found that 21% of the books randomly selected for rights clearance in the period 1870-2010 were in-commerce. However if the pre-1930s books that were available commercially entirely as print-on-demand were excluded, as being public domain in the United States²¹, this figure drops to just 11%. Further empirical evidence is required to establish whether this figure would hold water more broadly, but it is certainly a fact of the publishing industry that relatively few books make it beyond even a second or third reprint so it would perhaps not be inaccurate to assume that the overwhelming majority of in-copyright items from the 20th century that sit in the world's memory institutions are returning no royalties to their authors at all. This statement, although perhaps provocative to some, is also backed up by a study²² of US sound recordings commissioned by Congress that found that only 2% of sound recordings between 55 and 75 years old return royalties to their creators.

Since November 2010 a number of meetings have been convened by DG Internal Market mainly in Brussels to discuss the issue of rights clearance of in-copyright but out-of-commerce works in the context of mass digitisation by libraries, universities and other cultural bodies as defined by Art 5.2.(c) of the Information Society Directive. While the initial point of discussion was broad in scope, the majority of those convened for the meeting from the rightsholders' side were from the print world, and as the newspaper industry representatives distanced themselves from the discussions relatively early on, the stakeholder dialogue focussed mainly on printed books and journals. The stakeholders represented in the debate were from a varied assortment of sectors varying from national libraries and Europeana to research and public library associations, as well as various global, European and national publishers' associations, collecting societies, author groups and representatives of visual artists.

Over the course of a number of meetings between November 2010 and July 2011 a memorandum of understanding was arrived at that envisages a rights clearance mechanism for out-of-commerce works akin to those currently in operation in Scandinavia, namely one that recognises that in the context of mass digitisation, item by item clearance is impossible and therefore collecting societies have an important role to play in granting permission. The MOU, which was signed by all the stakeholders in the presence of the current Commissioner for DG Internal Market, Michel Barnier, recognises that in order to legally clear rights, there has to be a change in law in EU member states to allow collecting societies that are sufficiently

²¹ Pre 1923 creative works are treated in the US as public domain.

²² Rappaport, E. 1998. *Copyright Term Extension: Estimating the Economic Values*. Congressional Research Service. The Library of Congress.

representative of a particular category of creative work to extend their remit to not only artists and creators who are registered with them, but to non-members who produce works in the same category also. As this “extension effect” is essentially a limitation in law on the exclusive rights of a creator to decide how his or her work is used, the MOU recognises that the “extension effect” must be legal across European borders in order to ensure one member state’s solution to represent this class of creator will be recognised as legal in another. This is important as the ARROW report (Stratton, 2011) shows a large proportion of what is collected by a national library is grey literature - material not produced by commercial publishers but by many varying types of individual or institution for many different primarily non-commercial purposes. While no more than an MOU at this stage, its very existence combined with the key stakeholders involved in its preparation, evidently sends a strong message to the Commission as to the need for a next wave of legislation required to make mass digitisation in Europe a reality.

Data Analytics and the Granularisation of Knowledge

Just as the digital revolution has had a lasting impact on how we choose to purchase or disseminate copyright goods, it has also fundamentally affected how we are able to or desire to reuse knowledge itself. Copyright law as an analogue concept is based upon the concept of needing to regulate the copying of copyright goods in order to prevent the creation of rivalrous or substitutable products. Clearly the availability of copies of substitutable music tracks from both legal sources such as iTunes, as well as illegal sources that exist on the web, have the potential to undermine the sale of the original work. Similarly illegally copied books for sale in street markets will compete with the sale of the genuine work in a bookshop, or online from a website like Amazon.

The concept of copying, and the ability to control onward copying for the reasons outlined above lies at the heart of copyright law. This is illustrated by the fact that in UK copyright law (Copyright, Designs and Patents Act 1988), the first action that is defined as the exclusive right of the rightsholder to control is “to copy the work”. This principle clearly makes sense in an analogue environment where what could be done with a copyright work was relatively limited, either by the restrictions of the format itself, or to entities (usually companies) with the appropriate level of resources at their disposal to recast and copy the original material. In the digital world however copying is ubiquitous, far less costly, and varied in the form it can take. Also unlike in the analogue world, digital technology does not limit how a copyright work can be used, and therefore non-rivalrous works that are not substitutable for the original can be made easily - from sampling music, humorous internet memes, the Google translate service through to the use of a low resolution image copied and reworked for a Facebook page. It can be argued that such digital copies, unlike most analogue copies, do not compete with the original as they are by definition different in composition and form and are therefore non-substitutable or non-rivalrous.

In essence computer technology allows copying to happen in many different ways, less linear and replicate of the original and yet still having immense value. One very important example of this in the field of science is the technique of text and data mining. The process is a computer based one, that uses automated tools, often based on semantic technology, to search for trends and relationships by extracting information from a source, to establish a theory or hypotheses for further evaluation. This technique, nearly ten years old, is being increasingly used by scientists (but also researchers in the field of arts and humanities) to make new discoveries and improve the speed at which scientific innovation occurs. The technique is also widely used by finance companies, marketing companies, and journalists²³ to find information in a world of data and information plenty – the world of ‘big data’.

Given the plethora of digital information available, society has probably reached a tipping point where effective interaction with large volumes of information will have to rely on computers to aid discovery. We can see this already in the central role of search tools in all electronic platforms employed by search engines and online scholarly publications. Data analytics however is not

²³ <http://www.guardian.co.uk/technology/blog/2011/jun/23/knight-foundation-scrapewiki-okf-data> [Accessed 14/08/ 2011]

'search' in the sense employed by search engines or existing scientific platforms where a key word is simply sought. It employs techniques derived from the field of computational linguistics using a specific pre-programmed algorithm to extract from selected data or texts relevant information that expresses a relationship between entities or facts. The National Centre for Text Mining, based at the University of Manchester, defines the mining of text as "the process of discovering and extracting knowledge from unstructured data." Step one involves the identification of material to be analysed, and then crucially the standardisation or normalisation of the data (i.e. copying) into one standard format that can be interrogated by the relevant pre-programmed computer algorithm. Step two involves the extraction of the facts and relationships that are the subject of the research work, to create a derived dataset. Step three is the interrogation of the derived dataset to establish whether there is a relationship between the extracted facts or not.

The technique, still in its early days, however has already started to show immense potential, particularly in biomedical research. For example, scientists by mining the literature alone, and never stepping foot in a laboratory, have discovered links between therapies and drugs that have never before been linked in literature (Sainani, 2008). For example, the link between magnesium and migraines, Viagra and Parkinson's disease, fish oil and a disease of the blood vessels known as Raynaud's disease, or a new use of the thalidomide drug.²⁴ Other interesting medical discoveries include identifying particular genes that can be the target for gene therapy, important in the development of osteoporosis,²⁵ or genes that are important in how one's body responds to radiation.²⁶ It is not an exaggeration to suggest that this new computer technique has life-changing potential in the field of medical research and heralds with it the chance of speeding up the innovation cycle for the benefit of all mankind.

Given this, the imperative for ensuring that data analytics is exempted from the exclusive rights granted by copyright law would appear to be clear. It is however only starting to be discussed at a policy level in the West as part of the Hargreaves Review in the UK, with Japan being the only country in the world to have a specific limitation and exception for data analytics, which it implemented 2009.

The heart of this issue from the perspective of an intellectual policy law maker is one facet of a larger issue that faces copyright law itself in the digital age, namely that copying is now absolutely ubiquitous, woven into the fabric of all technology, and yet copyright law itself essentially is a prohibition of copying by anyone except the rightsholder. Not only this but as a result of the digital revolution the very use of copyright goods, which have in part because of the way analogue technologies work, never been subject to copyright law before have now become so because of the way technology works. From the forwarding of an email to the caching or accessing of a website to read, all is subject to copyright law by dint of the fact that a copy has been made on a computer in the process of performing a particular usage act. This contrasts with analogue uses of copyright works like sending a letter, reading, or listening to music on a record player as no copies are made, and therefore the acts of access and consumption of copyright material is unregulated. The fact that copyright now can cover the simple usage of a copyright work also represents a seismic shift, and to some a concerning shift, in what copyright law itself covers as traditionally only the intentional act of copying has been the restricted act, while 'use' has been free from control of the law.

It is for reasons such as this that it is argued by copyright scholars that copyright law has grown out of all recognition to what it was originally intended to do - namely be a policy tool, as the Statute of Anne states for the "encouragement of learning". Examples of how the scope of copyright has spread include the dramatic increase in duration of term from 14 years to now on

²⁴ Weeber et al. 2003,10. *Journal of the American Medical Informatics Association*. p.252–259

²⁵ Maté, O and L Dehaspe. 2010, 11(Suppl 5). *Integrating automated literature searches and text mining in biomarker discovery*. BMC Bioinformatics.

²⁶ Smirnov, D.A and M Morley, E Shin, RS Spielman, VG Cheung. 2009. *Genetic analysis of radiation-induced changes in human gene expression*. Nature.

average over 100 years, new intellectual property laws to protect the selection and arrangement of a database as a separate right to the content, the interpretation by an actor of say a play by Shakespeare when fixed in a recording, or patents that protect business methods. It is this, the expansion in length and scope of copyright and related rights that have led some scholars, to refer to this phenomenon as the “commodification of information” (Elkin-Koren, 2002). Elkin-Koren argues that what we are witnessing is “the turning of information into a perfect commodity, granting rightsholders a set of powerful legal rights to control every access and use of such information.”²⁷ The point being of course that the ability to, or space where, individuals can freely innovate, share and build on other peoples’ work has been diminished.

The area of data analytics is one such area where the public interest, and indeed the interest of certain industrial sectors are being curtailed because of the ever expanding ways that intellectual property laws are now framed. From an information law perspective, data analytics is on the fault line between copyright law and contract law. Whereas facts themselves are not subject to copyright, copyright law being designed to protect artistic expression, the way to extract them as outlined above involves copying of the entire work in order to standardise the data, prior to extracting the required information. Simply because computers must copy articles, texts or book chapters before the facts in the data can be extracted as part and parcel of text and data mining means that this process is subject to copyright law (as well as database law in the EU). Adding to the legal complexity is the fact that of course most electronic material is not covered by copyright but by contract law meaning that only if the copying and extraction of text is expressly permitted in the contract can text and data mining be performed lawfully. This double bind of copyright law and contract law creates a situation whereby even the use of facts, the very building blocks of knowledge, can be off limits to users – the perfect commodity as argued by Elkin-Koren.

It is with this background that researchers, pharmaceutical companies and technology companies have begun to argue against the absolute commodification of information. Perhaps it is not surprising that the first country in the world to adopt an exception for data analytics is Japan, a country well known for its cutting edge and innovative approach to technology. As part of the Japanese government’s preparation for amending Japanese copyright law in 2009, a report by the Copyright Subcommittee of the Department of Culture argued “In an advanced information society amidst vast volumes of information, data analytics technology, which allows the extraction of information as well as the advanced processing of such knowledge, is a necessity for users, as well as a fundamental of a digitally networked society. It can also be argued that the development of research involving data analytics has many societal benefits. In addition, another side to the argument is that research developments using data analytics do not use the (artistic) expression contained in a copyright work itself, as it is no more than the extraction of information. And that while in the process of data analytics a copyright work is used; its actual essence is not.”²⁸

The distinction between the unavoidable use of a copyright work almost as a by-product of technology, and the very real usage of artistic expression which copyright protects, is a theme further developed by the UK’s Intellectual Property Office (IPO) in its 2011 paper “Digital Opportunity – A Review of Intellectual Property and Growth”²⁹. It argues that stronger exceptions preventing contractual override are needed where the use of a copyright work does “not directly trade on the underlying creative and expressive purpose of the work.” In explaining its rationale for a new exception for data analytics the IPO elaborates by saying “the idea is to encompass the uses of copyright works where copying is only carried out as part of the way the technology works. For instance, in data mining or search engine indexing, copies need to be created for the computer to be able to analyse; the technology provides a substitute for someone reading all the documents. This is not about overriding the aim of copyright – these uses do not compete with the normal exploitation of the work itself – indeed, they may facilitate it. Nor is copyright intended to restrict uses of facts. That these new uses happen to fall within the scope of copyright

²⁷ Elkin-Koren. 2002. p84

²⁸ 2009. *Bunkashingikai Chosakukenbunkakai Hōkokusho*. Bunkashingikai Chosakukenbunkakai http://www.bunka.go.jp/chosakuken/pdf/21_houkaisei_houkokusho.pdf [Accessed 18/09/2011]

²⁹ <http://www.ipso.gov.uk/ipreview.htm>

regulation is essentially a side effect of how copyright has been defined, rather than being directly relevant to what copyright is supposed to protect.”

Such a debate is certainly timely, as a couple of recent cases in Europe demonstrate how the trend of commodification of information into smaller and smaller chunks of knowledge is developing. In these cases we see how control is being asserted over a few words and how information players legal resources are being deployed to keep a competitive lead in traditional information markets. One case that was taken to the European Court of Justice (ECJ) in 2009 between a Danish news monitoring service Infopaq, and an association of Danish daily newspaper publishers (Infopaq International A/S v Danske Dagblades Forening³⁰) revolved around whether the use of 11 words extracted from an article by the news monitoring company were subject to the exclusive rights of the rightsholder or not. The ECJ ruled conclusively that even such small amounts of information are indeed subject to copyright by stating that “an act occurring during a data capture process which consists of storing an extract of a protected work comprising 11 words and printing out that extract, is such as to come within the concept of reproduction ... if the elements thus reproduced are the expression of the intellectual creation of their author.”

Another more recent case is one from the UK – Newspaper Licensing Agency v Meltwater Holdings BV. This time involving the use of publicly available websites that Meltwater, another media monitoring organisation, ‘scrape’ to provide their clients with headlines, the url, the opening words of an article, and a number of words from the article surrounding pre-selected keywords. As in the Infopaq case, it was found that there was copyright in headlines or short extracts from newspaper articles, and therefore media monitoring is an entirely licensable activity. Some of the rulings that fall out of this case are of particular interest as they go to the heart of information sharing in a networked world and demonstrate how the smallest elements of knowledge are subject to legal battle for control. For example the High Court judge in the Meltwater case, referring to the Infopaq ruling, even rejected the need to examine on a case by case basis whether 11 words were original or creative acts and therefore attracted copyright, by arguing that as they contained the “tone of the article” which by definition would contain copyright, that the 11 words in the title would also be copyrightable. The ruling as it currently stands therefore creates a situation whereby effectively by no more than association to a larger body of work which undoubtedly will contain artistic expression, as few as 11 words will in all instances be subject to copyright law. The transcript of the appeal³¹ also is interesting in that in its summary of the original case it makes clear that in the opinion of the High Court the functionality of computers, the core action of copying, and simply the usage of this technology by definition leads users to infringe copyright. The transcript states “Accordingly the copies made by the end-user's computer of (a) Meltwater News (i) on receipt of the email from Meltwater, (ii) opening that email, (iii) accessing the Meltwater website by clicking on the link to the article and (b) of the article itself when (iv) clicking on the link indicated by Meltwater News are and each of them is, prima facie, an infringement of the Publishers' copyright.”

For libraries, and researchers, whose core function is the reuse of knowledge, this trend of granularisation and commodification of even the smallest shards of information raises many issues. For a library that catalogues analogue items, or scrapes information from websites that relate to its analogue or web collections, to scientists who wish to text and data mine facts, it could be argued that the information landscape in the digital era has been comprehensively fenced in. The consequences for research, education and innovation are potentially serious, given that knowledge and scientific information is held within works governed by contract law or intellectual property laws. The choice that faces citizens and policy makers is how the regulation of knowledge should develop in the future now the *modus operandi* of information regulation by a mutually reinforcing mixture of contract and copyright law in the digital world appears to have been set. There appear to be two choices: whether knowledge should be treated as a perfect commodity to be bought or sold, or whether the public interest in the unhindered flow of information supporting innovation and scientific discovery requires a reengagement by policy

³⁰ <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:220:0007:0007:EN:PDF>

³¹ <http://www.bailii.org/ew/cases/EWCA/Civ/2011/890.html> [Accessed 18/09/2011]

makers in limitations and exceptions. Without a reengagement with limitations and exceptions at a political level, it would certainly appear that the barriers to freely and lawfully reusing digital information have been placed very high indeed.

Electronic Legal Deposit

No discourse on modern copyright law and how it relates to library and researcher activity would be complete without some reference to legal deposit, and in particular the development of electronic legal deposit. Legal deposit developed in Europe in the 16th and 17th centuries as a legal mechanism to improve the development of a royal or national collection, as well as control certain trade privileges and monitor what was being published in the country at that time (Larivière, 2000). For example, the 1661 Swedish edict on legal deposit states that “it appears appropriate and useful that His Majesty would be informed of all the books and the writings printed in his kingdom.” In France, Napoleon also amended legal deposit law in 1810 to ensure that copies of published works were sent to the Ministry of Police to facilitate surveillance of the press.

Legal deposit became linked to copyright, and a required formality to enjoy legal protection for copyright in 1710 in the United Kingdom when the Statute of Anne was passed. The Act granted copyright for a period of 14 years, on the condition that nine copies of printed works were deposited in several libraries across the country. Between 1710 and the implementation of the Berne Convention in 1886, a number of countries including the United States, France and Belgium also went on to link the deposit of a publication in a legal deposit library to the enjoyment of copyright. Upon implementation of Berne, and the waiving of all requirements for copyright registration or ‘formalities’ as they are known, the two activities were decoupled even if legal deposit legislation may form part, particularly in a number of common law countries, of the copyright act to this day.

Naturally enough given the advent of digital publication technologies, and the perceived public interest in preserving a nation’s cultural and intellectual heritage many countries have started to draft electronic legal deposit legislation that deals with the deposit and acquisition of the national imprint in electronic form. The approach and the speed at which this has been done varies very much from country to country.

A 2009 examination by the British Library of 26 countries (all European except Australia, Japan, Canada, New Zealand and the United States) found that 74% of the countries surveyed had already passed and implemented a legal deposit obligation for at least one type of electronic material to be deposited with the national library. In the majority of cases this was an offline electronic publication like a CD-ROM. 54% of the countries did however report that they were able to lawfully harvest online freely available websites. Libraries from 10 countries also reported that electronic deposit legislation covered e-books, e-journals and other commercial or protected online publications. A further 6 libraries (Denmark, Finland, France, Germany, New Zealand and Norway) confirmed that existing legislation also includes an obligation to deposit structured datasets that sit behind query-driven services such as journey planners or telephone enquiry datasets.

Clearly different countries will approach the issue of electronic legal deposit differently influenced by factors ranging from their own traditions around this particular body of legislation to the pressures of the legislative timetable. It is however perhaps interesting, but not surprising to note that those countries that have made the quickest progress on legal deposit, are not the common law countries like the UK, US and Australia who have traditionally linked legal deposit to copyright law, but the civil law countries. That is to say countries who have generally not sought historically to incorporate or link the development of the national collection to copyright, but rather have seen legal deposit as having a primarily cultural role independent of the economic ethos of copyright that exists in common law.

Regulation of the Internet v Access to Knowledge

The imaginings of a decentralised network, with no governmental oversight, where people were able to exist, innovate and communicate freely with similarly minded people has already receded into the legend of how the internet was originally conceived and developed. The network has its early origins in 1969 with the connection of the first nodes of the ARPANET (Advanced Research Projects Agency Network) and grew exponentially until its 'going public' in 1994. In its early days it was predominantly developed and used by researchers, students, and technologists working for universities and large computing corporations. In literature on the topic the ethos that governed the earliest years of the internet is often referred to as "anarchic" or "libertarian". Time Magazine³² writing in 1994, described the basics of the early philosophy that dominated the internet to be a "hacker ethic". Quoting Wired writer Steven Levy the magazine summarised the main philosophical tenets of this group of early internet pioneers to be:

1. Access to computers should be unlimited and total.
2. All information should be free.
3. Mistrust authority and promote decentralisation.

Despite this the reality was that regulation of the web began almost immediately it began to be adopted outside of the limited confines of the research environment. In China for example, as early as 1994 the Ordinance for Security Protection of Computer Information Systems gave responsibility to the police for internet security protection to "supervise, inspect and guide the security protection work" and to "investigate and prosecute illegal criminal cases." In the United States in the same year surveillance legislation was also implemented when telecommunications companies were by law made to cooperate with law enforcement agencies in the form of the Communications Assistance for Law Enforcement Act (CALEA). In spite of the aspirations held by those who first developed the network the fact is surveillance of the internet of course is now not uncommon. For example Freedom of the Net 2011³³ in a survey of 37 countries from all six continents found that only eight countries' internets surveyed were "free", 18 "partly free" and 11 "not free."

In the liberal democracies of the developed world, the first widely referred to case of online governmental intervention that impacted upon libraries came in the aftermath of September 11th and the passing of the USA PATRIOT Act in 2001. The Act increased the right of US law enforcement agencies to examine records of information use by individuals without a warrant. A survey undertaken by the University of Illinois in 2002 found that in the year since September 11th 2001, 545 libraries out of 1505 surveyed had been approached by US law enforcement agencies for information about their patrons' reading habits as well as internet preferences (Kranich, 2003). Cases such as this, widespread internet monitoring particularly in the Middle East and Asia, as well as a healthy debate within the library community about voluntary library filtering of societally unacceptable web content led in 2002 to the International Federation of Library Associations (IFLA) publishing the *IFLA Internet Manifesto*. The manifesto asserts that intellectual freedom is closely intertwined with the freedom of access to information, and that protecting these principles "regardless of medium" is one of the core roles of the library community.

More recently in the past three to four years, in an attempt to combat the problem of online piracy being experienced by the entertainment industry, there has been an increasing global legislative focus on online copyright infringement. At a global level this has taken the form of the Anti-Counterfeiting Trade Agreement (ACTA)³⁴, the final text of which was issued in November 2010. A plurilateral agreement in the area of IPR enforcement, expected to be signed by the EU, US, Japan, South Korea, Canada, New Zealand, Singapore, Morocco, Australia and Switzerland, ACTA has proved controversial not least because despite it establishing an international legal framework concerning IPR enforcement, it was created outside of the existing international

³²1994. *Battle for the Soul of the Internet*. Time Magazine. <http://www.time.com/time/magazine/article/0,9171,981132,00.html> [Accessed 18/09/2011]

³³2011. *Freedom on the Net. A Global Assessment of Internet and Digital Media*. Freedom House. <http://www.freedomhouse.org/uploads/fotn/2011/FOTN2011.pdf> [Accessed 18/09/2011]

³⁴http://trade.ec.europa.eu/doclib/docs/2011/may/tradoc_147937.pdf

structures like the World Trade Organisation and World Intellectual Property Organisation. The scope of the agreement covers online copyright infringement, counterfeit goods and generic medicines. In regards to the internet, critics argue that it encourages the adoption of further web monitoring in the form of “graduated response” legislation which will inevitably lead to the disconnection of people from the internet.³⁵

This type of legislation known interchangeably as ‘three strikes’ ‘graduated response’, or in France the ‘Hadopi laws’ (after the acronym of the governmental body administering the legislation) has been implemented already in a number of countries – namely France, New Zealand, UK and South Korea – all of which are highly likely to be ACTA signatories. The intent of such legislation is to minimise online copyright infringement by requiring Internet Service Providers and rightsholder organisations to cooperate in order to identify individuals who have illegally shared or used copyright works unlawfully. After an alleged infringer has been written to three times regarding a suspected copyright infringement, they will be potentially disconnected from the internet or face other legal penalties.

Laws of this nature have proved controversial for many different reasons, ranging from them ignoring fundamental principles of law such as a presumption of innocence, through to how such legislation sits alongside other bodies of law, in particular privacy and data protection laws. The public debates have very much focussed on the ramifications for the main intended protagonists – namely ISPs, who have until now enjoyed clear and strong safe harbour provisions in law as “mere conduits”, and citizens who may be wrongly accused, and have a fundamental right to privacy. What the laws however have failed to understand is that the internet is more than a simple binary relationship of an individual householder subscribing directly to bandwidth from a commercial ISP. That is to say ‘the middle of the network’, where many differing types of organisations sit, has been simply ignored or more likely just not understood when drafting the legislation. Organisations as varied as schools, public libraries, universities, coffee shops, rail companies, hotels, colleges, local and even national governments all sit in the ‘middle’ providing internet access to individuals, but are not the commercial telecommunication providers or the individual subscribers that the law has been crafted around.

Despite the levels of press coverage around graduated response, with the notable exception of New Zealand the unintended consequences of the legislation for libraries has not been publicly explored in much detail. The main issue for libraries being that the detailed legislation that has been drawn up describes and establishes a simple rightsholder/ ISP/ individual consumer relationship, and interpreting the legislative position of libraries that sit across this neat and clean linear relationship has proven extremely problematic. In New Zealand, the Copyright (Infringing File Sharing) Amendment Bill came into force in September 2011, presenting challenges for public libraries as well as those that are part of universities, schools and colleges. The New Zealand Library Association (LIANZA) was reported widely criticising the Act as it is the account holder and not the individual who has infringed who is responsible for downloads on a connection. Because of this libraries in New Zealand face fines of up to \$15,000. In an article in the New Zealand Herald, a library spokesman for LIANZA also touched upon the likely consequences of being embroiled in infringement cases with risk averse local authorities demanding action to prevent further infringement potentially leading to calls by them to withdraw access in libraries to the internet altogether.³⁶ The issue facing libraries everywhere where three strikes legislation has been implemented is that many users use computers in public and university libraries and therefore identifying an individual can be impossible. In addition to the numbers of users on a single computer, it is not an uncommon practice for public bodies like libraries or universities to distribute a single IP address amongst a whole bank of computers again making identification of an infringing individual complex if not impossible.

³⁵ See letter from over 75 US law professors to President Obama on ACTA. http://www.wcl.american.edu/pijip/go/blog-post/academic-sign-on-letter-to-obama-on-acta#_ftn11 [Accessed 18/09/2011]

³⁶ Fletcher, H. 18/5/2011. *Library Internet Faces Axe*. The New Zealand Herald. http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10726286 [Accessed 18/09/2011]

The UK's version of three strikes legislation is known as the Digital Economy Act (DEA) which was passed in 2010, during the peculiarly British 'wash-up' process shortly before a general election. Not only was the impact of the Act on public intermediaries like libraries, schools and colleges not part of the regulatory impact assessment, the very core definitions of the Act make little sense when applied to any other entity than an individual householder or a commercial ISP. The core definitions of the Act defines an ISP as an entity that provides internet access to a subscriber, and provides "the allocation of an IP address or IP addresses to the subscriber to enable that access." The allocation of an IP address to a computer is a routine exercise administered by many libraries, and therefore as the Act is defined many libraries will be categorised as an ISP. On the other hand a subscriber is an entity who receives access to the internet "under an agreement" from the service provider. Again libraries contract for bandwidth and wi-fi "under an agreement" from ISPs, and therefore will simultaneously be classified as a subscriber. Clearly the presence of non-mutually exclusive definitions in an Act that places very distinct obligations and penalties on ISPs and subscribers is a legal blunder that is now unfortunately hard-baked into the DEA, leaving libraries in a situation of legal uncertainty as how the provisions of the Act actually apply to them. Whether libraries will be disconnected from access to the web because of actions by their users, or whether filtering and monitoring software will be increasingly adopted in public spaces in the UK to avoid being disconnected, or a mixture of the two, has yet to be seen. One thing that is certain though is that despite some levels of copyright infringement undoubtedly taking place across public networks, three strikes measures are a disproportionate measure when applied to libraries given their important public policy function.

In France as in the UK, the Hadopi laws passed in 2009 also do not specifically mention libraries. Hadopi, is an acronym of *Haute Autorité pour la diffusion des œuvres et la protection des droits sur internet*, the government agency who administers the three strikes law. In France, upon receiving information of a suspected online copyright infringement, the governmental body writes to the alleged infringer who is invited to install filtering software on their internet connection. As a result of a decree passed in June 2010 by the French government, it has also become an act of gross negligence not to install security software if requested to do so. The implementation will form a strong defence for an organisation like a library if infringements have been found to have taken place on their network, and therefore it seems likely that in order to remain compliant with the law, French libraries will adopt when requested, filtering and monitoring software on the computers they control.

For a library whose function is the provision of information not only is the prospect of being disconnected from the internet one that strikes at the heart of the *raison d'être* of the profession, but the prospect of filtering and monitoring is also one that raises perhaps equally complex issues. If the library profession is said to have a philosophical underpinning then it is probably Art.19 of the Universal Declaration of Human Rights which states that everyone has the right to "seek, receive and impart information and ideas through any media and regardless of frontiers." As a profession that is dedicated to providing access to information, whether in the traditional way by opening one's doors to one's shelved collections, or now through giving access to online electronic resources, librarians view themselves as neutral conduits to allow others to access information. Such a concept as freedom of access to information, irrespective of a librarian's own personal views, expresses itself in many ways from the production of guidance sheets in libraries on user privacy, opposition to the USA PATRIOT Act, through to monitoring and opposing censorship online and offline. From this perspective, 'three strikes' legislation that raises the spectre of either disconnection from the web, fines, or increased monitoring of individuals' online reading habits, is something that sits extremely uncomfortably with the underlying ethos and beliefs of a profession dedicated to information provision.

Conclusion

While traditional copyright law presents many intricate issues that affect how and when information is accessed, it is not absolutist. Limitations and exceptions allow access while a creative work is subject to the exclusive right of the rightsholder, and when copyright expires, the means to access the information at this point is not controlled by anyone. When a book or a sound recording has entered the public domain, as long as one has the physical ability to hear or view the item, or one is able to afford the means to reproduce them, access and reuse is always possible. This however cannot be said to be the same in the digital world.

Access to knowledge can now be denied in perpetuity by the copyright holder through the application of technical protection measures, or more commonly contract law. While certainly not the intent of the legislator, the slow or non-existent updating of preservation exceptions in certain countries may also mean that digital material will simply not exist, and citizens will not have access to their digital cultural heritage. Digital has also meant that the means to access knowledge itself are no longer either free of control. Whereas copyright law does not stop someone listening to a film, or taking notes from a book, in the digital world the copyright work as an object, and the means to access that object, are much more closely intertwined. In order to view a website or a purchased electronic database, internet access is of course a prerequisite. In order to read or use an electronic object a copy must be made. If three strikes legislation does lead to libraries being disconnected from web access, knowledge itself will no longer be accessible to the user as the very means of access have been removed. Similarly, while copyright law does not seek to control the fundamentals of knowledge sharing – namely facts and information – the interplay of contracts and computer functionality means that users are being locked out from the fundamental building blocks of learning and societal progress.

Access to information and knowledge is a fundamental of a sustainable society in the way that health, safety, shelter and food is. It is for this reason that we require an urgent re-evaluation of how we are seeking to regulate access to knowledge in the digital era. Without this innovation, scientific and cultural advancement will not be able to reach its full potential. Despite the life-changing possibilities digital has brought us, many copyright academics argue that we must seek to return to the underlying principles of copyright law and the norms of the analogue world. A world where the actual physical act of access and use of information could not be controlled, and a world where the public interest meant that for the benefit and welfare of all, government intervened to ensure the rules that governed access to knowledge were well calibrated, balanced and equitable. In the field of information law at least, perhaps the words of the French poet Alphonse de Lamartine should be viewed as a signpost for policy makers - “History teaches everything, even the future.” For without such a re-evaluation or “back to basics” approach by governments the future of knowledge and information sharing, crucial in an information society, seems to be an unpredictable and complex one fraught with uncertainties.

References

- Byrne, A. 2007. *The Politics of Promoting Freedom of Information and Expression in International Librarianship: The IFLA/FAIFE Project*. Scarecrow Press.
- Crews, K. 2008. *Study on Copyright Limitations and Exceptions for Libraries and Archives*. Standing Committee on Copyright and Related Rights. World Intellectual Property Organisation.
- Elkin-Koren, N. and N.W Netanel. 2002. *The Commodification of Information Law*. Kluwer Law International.
- Gibby R and A Stephens. *National Implementation of Electronic Legal Deposit*. (unpublished)
- Hamilton, S. 2004. *To what extent can libraries ensure free, equal and unhampered access to Internet-accessible information resources from a global perspective?* [Online] <http://archive.ifla.org/faife/report/StuartHamiltonPhD.pdf> (accessed 9/11/2011).
- Hilty, R and S Krujatz, B Bajon, A Früh, A Kur, J Drexl, C Geiger, N Klass. 2008. *European Commission – Green Paper: Copyright in the Knowledge Economy. Comments by the Max Planck Institute for Intellectual Property, Competition and Tax Law*.
- Kranich N. 2003. *The Impact of the USA PATRIOT Act on Free Expression*. [Online] <http://www.fepproject.org/commentaries/patriotact.html> (accessed 9/11/2011).
- Krikorian, G and A Kapczynski. 2010. *Access to Knowledge in the Age of Intellectual Property*. Zone Books.
- Lang, B. 2010/11. *Orphan Works and the Google Book Search Settlement: An International Perspective*. New York Law School Law Review. Volume 55.
- Larivière, J. 2000. *Guidelines for Legal Deposit Legislation*. International Federation of Library Associations and Institutions. [Online] <http://archive.ifla.org/VII/s1/gnl/legaldep1.htm> (accessed 18/09/2011)
- Ryan, J. 2010. *A History of the Internet and the Digital Future*. Reaktion Books.
- Sainani, K. 2008. *Mining Biomedical Literature: Using computers to extract knowledge nuggets*. Biomedical Computation Review.
- Stratton, B. 2011. *Seeking New Landscapes: A rights clearance study in the context of mass digitisation of 140 books published between 1870 and 2010*. British Library / ARROW.
- Vuopala, A. 2010. *Assessment of the orphan works issue and costs for rights clearance*. European Commission DG Information Society and Media Unit E4 Access to Information.
- Xalabarder, R. 2009. *Study on Copyright Limitations and Exceptions for educational activities in North America, Europe, Caucasus, Central Asia and Israel*. Standing Committee on Copyright and Related Rights. World Intellectual Property Organisation.